

From quantum Bayesian inference to quantum tomography *

R. Derka¹, V. Bužek^{1,2}, G. Adam³, and P.L. Knight²

¹ *Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, 842 28 Bratislava, Slovakia*

² *Optics Section, The Blackett Laboratory, Imperial College, London SW7 2BZ, England*

³ *Institut für Theoretische Physik, Technische Universität Wien,*

Wiedner Hauptstrasse 8-10, A-1040 Vienna, Austria

(December 13, 1996)

We derive an expression for a density operator estimated via Bayesian quantum inference in the limit of an infinite number of measurements. This expression is derived under the assumption that the reconstructed system is in a pure state. In this case the estimation corresponds to an averaging over a generalized microcanonical ensemble of pure states satisfying a set of constraints imposed by the measured mean values of the observables under consideration. We show that via the “purification” ansatz, statistical mixtures can also be consistently reconstructed via the quantum Bayesian inference scheme. In this case the estimation corresponds to averaging over the generalized canonical ensemble of states satisfying the given constraints, and the reconstructed density operator maximizes the von Neumann entropy (i.e., this density operator is equal to the generalized canonical density operator which follows from the Jaynes principle of maximum entropy). We study in detail the reconstruction of the spin-1/2 density operator and discuss the logical connection between the three reconstruction schemes, i.e., **(1)** quantum Bayesian inference, **(2)** reconstruction via the Jaynes principle of maximum entropy, and **(3)** discrete quantum tomography.

I. INTRODUCTION

The essence of the problem of state determination (\equiv state reconstruction) lies in an *a posteriori* estimation of a density operator (\equiv corresponding quasiprobability density distribution) of a quantum-mechanical (microscopic) system based on data obtained with the help of a macroscopic measurement apparatus [1]. The quality of the reconstruction depends on the quality of the measured data and the efficiency of the reconstruction procedure with the help of which the inversion-data analysis is performed. In particular, we can specify three different situations: Firstly, when all system observables are precisely measured; Secondly, when just part of the system observables is precisely measured; Finally, when measurement does not provide us with information enough to specify exact mean values (or probability distributions) of observables under consideration.

A. Complete observation level

Providing all system observables (i.e., the quorum [2]) have been precisely measured, then the density operator of a quantum-mechanical system can be completely reconstructed (i.e., the density operator can be uniquely determined based on the available data). In principle, we can consider two different schemes for reconstruction of the density operator (or, equivalently, the Wigner function) of the given quantum-mechanical system. The difference between these two schemes is based on the way in which information about the quantum-mechanical system is obtained. The first type of measurement is such that on each element of the ensemble of the measured states only a *single* observable is measured. In the second type of measurement a *simultaneous* measurement of conjugated observables is assumed. We note that in both cases we assume an ideal, i.e., unit-efficiency, measurements.

*We dedicate this paper to the sixtieth birthday of Professor Jan Peřina.

1. Quantum tomography

When the single-observable measurement is performed, a *distribution* $W_{|\Psi\rangle}(A)$ for a particular observable \hat{A} in the state $|\psi\rangle$ is obtained in an unbiased way [3], i.e., $W_{|\Psi\rangle}(A) = |\langle\Phi_A|\Psi\rangle|^2$, where $|\Phi_A\rangle$ are eigenstates of the observable \hat{A} such that $\sum_A |\Phi_A\rangle\langle\Phi_A| = \hat{1}$. Here a question arises: What is the *smallest* number of distributions $W_{|\Psi\rangle}(A)$ required to determine the state uniquely? If we consider the reconstruction of the state of a harmonic oscillator, then this question is directly related to the so-called Pauli problem [4] of the reconstruction of the wave-function from distributions $W_{|\Psi\rangle}(q)$ and $W_{|\Psi\rangle}(p)$ for the position and momentum of the state $|\Psi\rangle$. As shown by Gale, Guth and Trammel [5] for example, the knowledge of $W_{|\Psi\rangle}(q)$ and $W_{|\Psi\rangle}(p)$ is not in general sufficient for a complete reconstruction of the wave (or, equivalently, the Wigner) function. In contrast, one can consider an *infinite* set of distributions $W_{|\Psi\rangle}(x_\theta)$ of the rotated quadrature $\hat{x}_\theta = \hat{q} \cos \theta + \hat{p} \sin \theta$. Each distribution $W_{|\Psi\rangle}(x_\theta)$ can be obtained in a measurement of a *single* observable \hat{x}_θ , in which case a detector (filter) is prepared in an eigenstate $|x_\theta\rangle$ of this observable. It has been shown by Vogel and Risken [6] that from an infinite set (in the case of the harmonic oscillator) of the measured distributions $W_{|\Psi\rangle}(x_\theta)$ for all values of θ such that $[0 < \theta \leq \pi]$, the Wigner function can be reconstructed uniquely via the inverse Radon transformation. In other words knowledge of the set of distributions $W_{|\Psi\rangle}(x_\theta)$ is equivalent to knowledge of the Wigner function (or, equivalently, the density operator). This scheme for reconstruction of the Wigner function (the so called *optical homodyne tomography*) has recently been realized experimentally by Raymer and his coworkers [7]. In these experiments Wigner functions of a coherent state and a squeezed vacuum state have been reconstructed from tomographic data. Quantum-state tomography can be applied not only to optical fields (harmonic oscillators) but for reconstruction of other physical systems, such as atomic waves (see recent work by Janicke and Wilkens [8]). Leonhardt [9] has recently developed a theory of quantum tomography of discrete Wigner functions describing states of quantum systems with finite-dimensional Hilbert spaces (i.e., angular momentum or spin).

2. Filtering with quantum rulers

In the case of the simultaneous measurement of two non-commuting observables (let us say \hat{q} and \hat{p}) it is not possible to construct a joint eigenstate of these two operators, and therefore it is inevitable that the simultaneous measurement of two non-commuting observables introduces additional noise (of quantum origin) into measured data. This noise is associated with Heisenberg's uncertainty relation and it results in a specific "smoothing" (equivalent to a reduction of resolution) of the original Wigner function of the system under consideration (see Refs. [10] and [11]). To describe a process of simultaneous measurement of two non-commuting observables, Wódkiewicz [12] has proposed a formalism based on an operational probability density distribution which explicitly takes into account the action of the measurement device modelled as a "filter" (quantum ruler). A particular choice of the state of the ruler samples a specific type of accessible information concerning the system, i.e., information about the system is biased by the filtering process. The quantum-mechanical noise induced by filtering formally results in a smoothing of the original Wigner function of the measured state [10,11], so that the operational probability density distribution can be expressed as a convolution of the original Wigner function and the Wigner function of the filter state. In particular, if the filter is considered to be in its vacuum state then the corresponding operational probability density distributions is equal to the Husimi (Q) function [10]. The Q function of optical fields has been experimentally measured using such an approach by Walker and Carroll [13]. The direct experimental measurement of the operational probability density distribution with the filter in an arbitrary state is feasible in an 8-port experimental setup of the type used by Noh, Fougères and Mandel [14] (for more details, see the recent book by Peřina and coworkers [15]).

As a consequence of a simultaneous measurement of non-commuting observables the measured distributions are fuzzy (i.e., they are equal to smoothed Wigner functions). Nevertheless, if the detectors used in the experiment have a unit efficiency (in the case of an ideal measurement), the noise induced by quantum filtering can be "separated" from the measured data and the density operator (Wigner function) of the measured system can be "extracted" from the operational probability density distribution. In particular, the Wigner function can be uniquely reconstructed from the Q function. This extraction procedure is technically quite involved and it suffers significantly if additional stochastic noise due to imperfect measurement is present in the data.

We note that propensities, and in particular Q -functions, can be also associated with discrete phase space and they can in principle be measured directly [16]. These discrete probability distributions contain complete information about density operators of measured systems.

B. Reduced observation levels and MaxEnt principle

As we have already indicated it is now well understood that density operators (or Wigner functions) can, in principle, be uniquely reconstructed using either the single observable measurements (optical homodyne tomography) or the simultaneous measurement of two non-commuting observables. The completely reconstructed density operator (or, equivalently, the Wigner function) contains information about *all* independent moments of the system operators. For example, in the case of the quantum harmonic oscillator, the knowledge of the Wigner function is equivalent to the knowledge of all moments $\langle (\hat{a}^\dagger)^m \hat{a}^n \rangle$ of the creation (\hat{a}^\dagger) and annihilation (\hat{a}) operators.

In many cases it turns out that the state of a harmonic oscillator is characterized by an *infinite* number of independent moments $\langle (\hat{a}^\dagger)^m \hat{a}^n \rangle$ (for all m and n). Analogously, the state of a quantum system in a finite-dimensional Hilbert space can be characterized by a very large number of independent parameters. A *complete* measurement of these moments can take an infinite time to perform. This means that even though the Wigner function can in principle be reconstructed the collection of a complete set of experimental data points is (in principle) a never ending process. In addition the data processing and numerical reconstruction of the Wigner function are time consuming. Therefore experimental realization of the reconstruction of the density operators (Wigner functions) for many systems can be problematic.

In practice, it is possible to perform a measurement of just a finite number of independent moments of the system operators which means that only a subset \hat{G}_ν ($\nu = 1, 2, \dots, n$) of observables from the quorum (this subset constitutes the so-called observation level [17]) is measured. In this case, when the complete information about the system is not available, one needs an additional criterion which would help to reconstruct (or estimate) the density operator uniquely. Provided mean values of all observables on the given observation level are measured precisely, then the density operator (the Wigner function) of the system under consideration can be reconstructed with the help of the Jaynes principle of maximum entropy (the so called *MaxEnt* principle) [17]. The reconstructed density operator fulfills several conditions. Firstly, its trace has to be equal to unity (i.e., $\text{Tr} \hat{\rho} = 1$). Secondly, $\text{Tr}(\hat{\rho} \hat{G}_\nu) = G_\nu$ ($\nu = 1, 2, \dots, n$) which means that the reconstructed density operator provides us with the measured mean values of those observables which constitute the given observation level. Obviously, a large number of density operators can fulfill these two constraints. So one needs an additional criterion which would uniquely specify the generalized canonical density operator. According to Jaynes [17] this operator has to be that one with the largest value of the von Neumann entropy $S = -\text{Tr}\{\hat{\rho} \ln \hat{\rho}\}$. This additional condition means that the *MaxEnt* principle is the most conservative assignment in the sense that it does not permit one to draw any conclusion not warranted by the experimental data [18].

The MaxEnt principle provides us with a very efficient prescription how to reconstruct density operators of quantum-mechanical systems providing mean values of a given set of observables are known. It works perfectly well for systems with (semi)infinite Hilbert spaces (such as quantum-mechanical harmonic oscillator) as well as for systems with finite-dimensional Hilbert spaces (such as spin systems). If the observation level is composed of the quorum of the observables (i.e. this is a complete observation level), then the MaxEnt principle represents an alternative to quantum tomography, i.e. both schemes are equally suitable for the analysis of the tomographic data (for details see [19]). To be specific, the observation level in this case is composed of all projectors associated with probability distributions of rotated quadratures. The power of the MaxEnt principle can be appreciated in analysis of incomplete tomographic data (equivalent to a reconstruction of the Wigner function in the discrete phase space). In particular, Wiedemann [20] has performed a numerical reconstruction of the density operator (Wigner function) from incomplete tomographic data based on the MaxEnt principle as discussed by Bužek et al. [19]. Wiedemann has shown that in particular cases *MaxEnt* reconstruction from incomplete tomographic data can be several orders better compared to a standard tomographic inversion. These results can be interpreted as suggesting that the MaxEnt principle is the conceptual basis behind the tomographic reconstruction (irrespective whether in continuous or discrete phase spaces).

C. Incomplete measurement and Bayesian inference

It has to be stressed that the Jaynes principle of maximum entropy can consistently be applied only when *exact* mean values of the measured observables are available. This condition implicitly assumes that an infinite number of repeated measurements on different elements of the ensemble has to be performed to reveal the exact mean value of the given observable. In practice only a finite number of measurements can be performed. What is obtained from these measurements is a specific set of data indicating number of how many times eigenvalues of given observables have appeared (which in the limit of an infinite number of measurements results in the corresponding quantum probability distributions). The question is, how to obtain the best *a posteriori* estimation of the density operator based on the measured data. Helstrom [21], Holevo [22], and Jones [23] have shown that the answer to this question can be given by the Bayesian inference method, providing it is *a priori* known that the quantum-mechanical state which is going

to be reconstructed is prepared in a pure (even though unknown) state. Once this purity condition is fulfilled, then the observer can systematically estimate (i.e. reconstruct) an *a posteriori* density operator based upon an incomplete set of experimental data. This density operator is equal to the mean over all possible *pure* states weighted by a specific probability distribution in an abstract state space with the unique invariant integration measure. It is this probability distribution (conditioned by the assumed Bayesian prior) which characterizes observer's knowledge of the system at every moment during the measurement sequence. We note once again that the Bayesian inference has been developed for a reconstruction of *pure* quantum mechanical states and in this sense it corresponds to an averaging over a *generalized microcanonical* ensemble.

In a real situation one can never design a state-preparation device such that it produces an ensemble of identical pure states. What usually happens is that the ensemble consists of a set of pure states, each of which is represented in the ensemble with a certain probability (alternatively, we can say that the system under consideration is entangled with other quantum-mechanical systems). So now the question is how to use the Bayesian reconstruction scheme when the quantum-mechanical system under consideration is in an impure state (i.e., a statistical mixture). To apply the Bayesian inference scheme, one has to define exactly three objects: **(1)** the abstract state space of the measured system; **(2)** the corresponding invariant integration measure of this space; and **(3)** the *prior* (i.e., the *a priori* known probability distribution on the given parametric state space). Once these objects are specified one can estimate an *a posteriori* density operator after each individual outcome of the measurement has been registered.

The main purpose of the present paper is to show from an example of the state-reconstruction of a spin-1/2 system how the Bayesian scheme of quantum inference developed for a reconstruction of statistical mixtures [24] actually works. We will show that this scheme corresponds to a specific averaging over the grand canonical ensemble. Moreover, we will show that in the limit of infinite number of measurements the reconstructed density operator is *equal* to the generalized canonical density operator obtained via the Jaynes principle of maximum entropy. In addition, in the case the complete observation level (the quorum of the observables is measured) the generalized canonical density operator is equal to the operator obtained via the tomographic measurement. This clearly reveals a logical connection between quantum Bayesian inference and quantum tomography.

The paper is organized as follows. In Section II we briefly review the Bayesian inference scheme as developed for pure states by Jones [23]. In Section III we derive limiting formula for an *a posteriori* estimated density operator of any quantum mechanical system under the prior assumption that the system is in a pure state. Section IV is devoted to a description of Bayesian inference for statistical mixtures. Reconstructions of a density operator of a spin-1/2 are presented in Section V. We will conclude our paper with general remarks.

II. BAYESIAN INFERENCE

The general idea of the Bayesian reconstruction scheme is based on manipulations with probability distributions in parametric state spaces. To understand this reconstruction scheme we remind us several definitions and concepts. Firstly, it is a space of states of the measured system. The quantum Bayesian method as discussed in the literature [21–23] is based on the assumption that the reconstructed system is in a pure state described by a state vector $|\Psi\rangle$, or equivalently by a pure-state density operator $\hat{\rho} = |\Psi\rangle\langle\Psi|$. The manifold of all pure states is a continuum which we denote as Ω . Secondly, it is the discrete space A of reading states of a measuring apparatus associated with the observable \hat{O} . These states are intrinsically related to the projectors $\hat{P}_{\lambda_i, \hat{O}}$, where λ_i are the eigenvalues of the observable \hat{O} .

The Bayesian reconstruction scheme is formulated as a three-step inversion procedure:

(1) As a result of a measurement a conditional probability

$$p(\hat{O}, \lambda_i | \hat{\rho}) = \text{Tr} \left(\hat{P}_{\lambda_i, \hat{O}} \hat{\rho} \right), \quad (2.1)$$

on the discrete space A is defined. This conditional probability distribution specifies a probability of finding the result λ_i if the measured system is in a particular state $\hat{\rho}$.

(2) To perform the second step of the inversion procedure one has to specify an *a priori* distribution $p_0(\hat{\rho})$ defined on the space Ω . This distribution describes our initial knowledge concerning the measured system. Using the conditional probability distribution $p(\hat{O}, \lambda_i | \hat{\rho})$ and the *a priori* distribution $P_0(\hat{\rho})$ we can define the *joint* probability distribution $p(\hat{O}, \lambda_i; \hat{\rho})$

$$p(\hat{O}, \lambda_i; \hat{\rho}) = p(\hat{O}, \lambda_i | \hat{\rho}) p_0(\hat{\rho}), \quad (2.2)$$

on the space $\Omega \otimes A$. We note that if no initial information about the measured system is known, then the prior $p_0(\hat{\rho})$ has to be assumed to be constant (this assumption is related to the Laplace principle of indifference [18]).

(3) The final step of the Bayesian reconstruction is based on the well known Bayes rule $p(x|y)p(y) = p(x; y) = p(y|x)p(x)$, with the help of which we find the conditional probability $p(\hat{\rho}|\hat{O}, \lambda_i)$ on the state space Ω :

$$p(\hat{\rho}|\hat{O}, \lambda_i) = \frac{p(\hat{O}, \lambda_i, \hat{\rho})}{\int_{\Omega} p(\hat{O}, \lambda_i, \hat{\rho}) d\Omega}, \quad (2.3)$$

from which the reconstructed density operator can be obtained [see Eq.(2.4)].

In the case of the repeated N -trial measurement, the reconstruction scheme consists of an iterative utilization of the three-step procedure as described above. After the N -th measurement we use as an input for the prior distribution the conditional probability distribution which is an output after the $(N - 1)$ st measurement. However, we can equivalently define the N -trial measurement conditional probability $p(\{ \} _N | \hat{\rho}) = \prod_{i=1}^N p(\hat{O}_i, \lambda_j | \hat{\rho})$ and applying the three-step procedure just once we get the reconstructed density operator

$$\hat{\rho}(\{ \} _N) = \frac{\int_{\Omega} p(\{ \} _N | \hat{\rho}) \hat{\rho} d\Omega}{\int_{\Omega} p(\{ \} _N | \hat{\rho}) d\Omega}, \quad (2.4)$$

where $\hat{\rho}$ in the r.h.s. of Eq.(2.4) is a properly parameterized density operator in the state space Ω . At this point we should mention one essential problem in the Bayesian reconstruction scheme, which is the determination of the integration measure $d\Omega$ ¹.

The integration measure has to be invariant under unitary transformations in the space Ω . This requirement uniquely determines the form of the measure. However, this is no longer valid when Ω is considered to be a space of mixed states formed by all convex combinations of elements of the original pure state space Ω . Although the Bayesian procedure itself does not require any special conditions imposed on the space Ω , the ambiguity in determination of the integration measure has been the main obstacle in generalization of the Bayesian inference scheme for a reconstruction of *a priori* impure quantum states.

III. BAYESIAN INFERENCE IN LIMIT OF INFINITE NUMBER OF MEASUREMENTS

The explicit evaluation of an *a posteriori* estimation of the density operator $\hat{\rho}(\{ \} _N)$ is significantly limited by technical difficulties when integration over parametric space is performed [see Eq.(2.4)]. Even for the simplest quantum systems and for a relatively small number of measurements, the reconstruction procedure can be technically insurmountable problem.

On the other hand let us assume that the number of measurements of observables \hat{O}_i approaches infinity (i.e. $N \rightarrow \infty$). It is clear that in this case mean values of all projectors $\langle \hat{P}_{\lambda_j, \hat{O}_i} \rangle$ associated with the observables \hat{O}_i are *precisely* know (measured): i.e.

$$\langle \hat{P}_{\lambda_j, \hat{O}_i} \rangle = \alpha_j^i, \quad (3.1)$$

where $\sum_j \alpha_j^i = 1$. In this case the integral in the right-hand side of Eq.(2.4) can be significantly simplified with the help of the following lemma:

Lemma: Let us define the integral expression

$$I(\alpha_1, \dots, \alpha_{n-1}) \equiv \int_0^1 dx_1 \int_0^{y_2} dx_2 \dots \int_0^{y_{n-1}} dx_{n-1} F(x_1, \dots, x_{n-1} | \alpha_1, \dots, \alpha_{n-1}) \quad (3.2)$$

where

$$F(x_1, \dots, x_{n-1} | \alpha_1, \dots, \alpha_{n-1}) = \frac{1}{B} x_1^{\alpha_1 N} x_2^{\alpha_2 N} \dots x_{n-1}^{\alpha_{n-1} N} (1 - x_1 \dots - x_{n-1})^{\alpha_n N} \quad (3.3)$$

and α_j satisfy condition $\sum_j^n \alpha_j = 1$. The integration boundaries y_k are given by relations:

¹ Many authors (see, for instance, Ref. [23]) identify the prior distribution with the integration measure on the space Ω . However, the particular form of $d\Omega$ is associated with the topology and the particular parameterization of the space Ω rather than with some prior information $p_0(\hat{\rho})$ about this system. We will distinguish between these two objects.

$$y_k = 1 - \sum_{j=1}^{k-1} x_j; \quad k = 2, \dots, n-1 \quad (3.4)$$

and B equals to the product of Beta functions $B(x, y)$:

$$B \equiv B(a_n + 1, a_{n-1} + 1) B(a_n + a_{n-1} + 1, a_{n-2} + 2) \dots B(a_n + a_{n-1} \dots a_2 + 1, a_1 + n - 1). \quad (3.5)$$

i. The function $F(x_1, \dots, x_{n-1} | \alpha_1, \dots, \alpha_{n-1})$ in the integral (3.2) is a normalized probability distribution in the $(n-1)$ -dimensional volume given by integration boundaries.

ii. For $N \rightarrow \infty$, this probability distribution has the following properties:

$$\langle x_i \rangle \rightarrow \alpha_i \quad \langle x_i^2 \rangle \rightarrow \alpha_i^2 \quad i = 1, 2, 3, \dots, n-1, \quad (3.6)$$

i.e., this probability density tends to the product of delta functions:

$$\lim_{N \rightarrow \infty} F(x_1, \dots, x_{n-1} | \alpha_1, \dots, \alpha_{n-1}) = \delta(x_1 - \alpha_1) \delta(x_2 - \alpha_2) \dots \delta(x_{n-1} - \alpha_{n-1}). \quad (3.7)$$

Proof: *i.* can be derived by a successive application of an equation [see for example [25], Eqs.(3.191)]

$$\int_0^u x^{\nu-1} (u-x)^{\mu-1} dx = u^{\mu+\nu-1} B(\mu, \nu). \quad (3.8)$$

ii. can be obtained as a result of a straightforward calculations of limits of certain expressions containing Beta functions with integer-number arguments. In our calculations we have used the identity

$$\frac{B(n+1, m)}{B(n, m)} = \frac{n}{n+m}, \quad (3.9)$$

which is satisfied by Beta functions with integer-number arguments.

A. Conditional density distribution

Let us start with the expression for conditional probability distribution $p(\{ \}_N | \hat{\rho})$ for the N -trial measurement of a set of observables \hat{O}_i . If we assume, that the number of measurements of each observable \hat{O}_i goes to the infinity then we can write:

$$p(\{ \}_{N \rightarrow \infty} | \hat{\rho}) = \lim_{N \rightarrow \infty} \prod_i \left[\prod_{j=1}^{n_i} \text{Tr} \left(\hat{P}_{\lambda_j, \hat{O}_i} \hat{\rho} \right)^{\alpha_j^i N} \right]. \quad (3.10)$$

The first product in the right-hand side (r.h.s.) of Eq.(3.10) is associated with each measured observable \hat{O}_i on a given observation level. The second product runs over eigenvalues n_i of each observable \hat{O}_i .

In what follows we formally rewrite the r.h.s. of Eq.(3.10): we insert in it a set of δ -function and we perform the following integration

$$\begin{aligned} p(\{ \}_{N \rightarrow \infty} | \hat{\rho}) &= \prod_i \left\{ \int_0^1 dx_1^i \int_0^{y_2^i} dx_2^i \dots \int_0^{y_{n_i-1}^i} dx_{n_i-1}^i \delta \left[x_1^i - \text{Tr} \left(\hat{P}_{\lambda_1, \hat{O}_i} \hat{\rho} \right) \right] \dots \right. \\ &\quad \left. \times \delta \left[x_{n_i-1}^i - \text{Tr} \left(\hat{P}_{\lambda_{n_i-1}, \hat{O}_i} \hat{\rho} \right) \right] \prod_{j=1}^{n_i-1} (x_j^i)^{\alpha_j^i N} (1 - x_1^i \dots x_{n_i-1}^i)^{\alpha_{n_i}^i N} \right\} \end{aligned} \quad (3.11)$$

In Eq.(3.11) we perform an integration over a volume determined by the integration boundaries y_k^i [see Eq.(3.4)], i.e., due to the condition $\sum_{j=1}^{n_i} \text{Tr}(\hat{P}_{\lambda_j, \hat{O}_i} \hat{\rho}) = 1$, there is no need to perform integration from $-\infty$ to ∞ .

At this point we utilize our *Lemma*. To be specific, firstly we separate in Eq.(3.11) the term, which corresponds to the function I given by Eq.(3.2). Then we replace this term by its limit expression (3.7). After a straightforward integration over variables x_j^i we finally obtain an explicit expression for the conditional probability $p(\{ \}_{N \rightarrow \infty} | \hat{\rho})$ which

we insert into Eq.(2.4) from which we obtain the expression for an *a posteriori* estimation of the density operator $\hat{\rho}(\{ \}_{N \rightarrow \infty})$ on the given observation level:

$$\hat{\rho}(\{ \}_{N \rightarrow \infty}) = \frac{1}{\mathcal{N}} \int_{\Omega} \prod_i \left\{ \prod_{j=1}^{n_i-1} \delta \left[\text{Tr} \left(\hat{P}_{\lambda_j, \hat{O}_i} \hat{\rho} \right) - \alpha_j^i \right] \right\} \hat{\rho} d\Omega. \quad (3.12)$$

Here \mathcal{N} is a normalization constant determined by the condition $\text{Tr} [\hat{\rho}(\{ \}_{N \rightarrow \infty})] = 1$.

The interpretation of Eq.(3.12) is straightforward. The reconstructed density operator is equal to the sum of equally-weighted pure-state density operators on the manifold Ω , which do satisfy the conditions given by Eq.(3.1) [these conditions are guaranteed by the presence of δ -functions in the r.h.s. of Eq.(3.12)]. In terms of statistical physics Eq.(3.12) can be interpreted as an averaging over the generalized microcanonical ensemble of those pure states which do satisfy the conditions on the mean values of the measured observables. Consequently, Eq.(3.12) represents the principle of the “maximum entropy” on the generalized microcanonical ensemble under the constraint (3.1).

In order to clarify the relationship between the reconstruction procedure based on the Jaynes principle of maximum entropy and the quantum Bayesian inference we present in Section V an example of the reconstruction of the state of a spin-1/2 system.

IV. BAYESIAN RECONSTRUCTION OF IMPURE STATES

In classical statistical physics a mixture state is interpreted as a statistical average over an ensemble in which any individual realizations is in a pure state. In quantum physics a mixture can be considered a state of a quantum system, which can not be completely described in terms of its own Hilbert space, because it is only a part of a more complex quantum system. Due to the lack of information about other parts of this complex system, the description of subsystem is possible only in terms of mixtures.

Let assume quantum system P being entangled with some other quantum system R (reservoir). Let the composed system S (composed of P and R) itself is in a pure state $|\Psi\rangle$. The density operator $\hat{\rho}_P$ of the system P is then obtained via tracing over the reservoir degrees of freedom:

$$\hat{\rho}_P = \text{Tr}_R [\hat{\rho}_S]; \quad \hat{\rho}_S = |\Psi\rangle\langle\Psi|. \quad (4.1)$$

Once the system S is in a pure state, then we can safely apply the Bayesian reconstruction scheme as described in Section 2. The reconstruction itself is based only on data associated with measurements performed on the system P . When the density operator $\hat{\rho}_S$ is *a posteriori* estimated, then by tracing over the reservoir degrees of freedom, we obtain the *a posteriori* density operator $\hat{\rho}_P$ for the system P (with no *a priori* constraint on the purity of the state of the system P). These arguments are equivalent to the “purification” Ansatz as proposed by Uhlmann [26].

To make our reconstruction scheme for impure state consistent we have to chose the reservoir R uniquely. This can be done with the help of the Schmidt theorem (see Ref. [27]) from which it follows that if the composite system S is in a pure state $|\Psi\rangle$ then its state vector can be written in the form:

$$|\Psi\rangle = \sum_{i=1}^M c_i |\alpha_i\rangle_P \otimes |\beta_i\rangle_R, \quad (4.2)$$

where $|\alpha_i\rangle_P$ and $|\beta_i\rangle_R$ are elements from two specific orthonormalized bases associated with the subsystems P and R , respectively, and c_i are appropriate complex numbers satisfying the normalization condition $\sum |c_i|^2 = 1$. The maximal index of summation (M) in Eq.(4.2) is given by the dimensionality of the Hilbert space of the system P . In other words, when we apply the Bayesian method to the case of impure states of M -level system, it is sufficient to “couple” this system to an M -dimensional “reservoir”. Due to the fact that we measure only observables of the first subsystem P particular form of states $|\beta_i\rangle_R$ of the second subsystem R does not affect results of the reconstruction.

V. SPIN-1/2 RECONSTRUCTION

We assume an ensemble of spins-1/2 in an unknown state described by the density operator

$$\hat{\rho}(\theta, \phi) = \frac{1}{2} \left(\hat{1} + \vec{r} \hat{\sigma} \right) \quad (5.1)$$

where $\vec{r} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$; $\phi \in (0, 2\pi)$, $\theta \in (0, \pi)$, and $\hat{1}$ is the unity operator. The Pauli spin operators $\vec{\sigma}$ in the matrix representation in the basis $|0\rangle, |1\rangle$ of the eigenvectors of the operator σ_z do read

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (5.2)$$

To determine completely the unknown state, one has to measure three linearly independent (e.g., orthogonal) projections of the spin-1/2. One possible choice of the complete set of observables (i.e., the quorum [2]) associated with the spin-1/2 are spin projections for three orthogonal directions represented by the Hermitian operators:

$$\hat{s}_i \equiv \frac{\hat{\sigma}_i}{2}, \quad i = x, y, z \quad (5.3)$$

In what follows we will consider three observation levels defined as $\mathcal{O}_A = \{\hat{s}_z\}$, $\mathcal{O}_B = \{\hat{s}_z, \hat{s}_x\}$ and $\mathcal{O}_C = \{\hat{s}_z, \hat{s}_x, \hat{s}_y\} \equiv \mathcal{O}_{comp}$.

In order to apply the Bayesian inference scheme for impure states, we have to “purify” [26] the quantum-mechanical system as discussed in Section IV. In the particular case of the spin-1/2 it means that we have to consider a system of two spins-1/2 on of which does play the role of a reservoir. In this case the corresponding two-spins-1/2 density operator can be parameterized as

$$\begin{aligned} \hat{\rho}(\alpha, \psi, \phi_1, \theta_1, \phi_2, \theta_2) = & \frac{\hat{1} \otimes \hat{1}}{4} + \frac{\vec{r}^{(1)} \hat{\sigma} \otimes \vec{r}^{(2)} \hat{\sigma}}{4} + \cos \alpha \left[\frac{\vec{r}^{(1)} \hat{\sigma} \otimes \hat{1}}{4} + \frac{\hat{1} \otimes \vec{r}^{(2)} \hat{\sigma}}{4} \right] \\ & + \sin \alpha \cos \psi \left[\frac{\vec{k}^{(1)} \hat{\sigma} \otimes \vec{k}^{(2)} \hat{\sigma}}{4} - \frac{\vec{l}^{(1)} \hat{\sigma} \otimes \vec{l}^{(2)} \hat{\sigma}}{4} \right] - \sin \alpha \sin \psi \left[\frac{\vec{k}^{(1)} \hat{\sigma} \otimes \vec{l}^{(2)} \hat{\sigma}}{4} + \frac{\vec{l}^{(1)} \hat{\sigma} \otimes \vec{k}^{(2)} \hat{\sigma}}{4} \right], \end{aligned} \quad (5.4)$$

where $\psi, \phi_1, \phi_2 \in (0, 2\pi)$; $\alpha, \theta_1, \theta_2 \in (0, \pi)$ and

$$\begin{aligned} \vec{k}^{(j)} &= (\sin \phi_j, -\cos \phi_j, 0); \\ \vec{l}^{(j)} &= (\cos \theta_j \cos \phi_j, \cos \theta_j \sin \phi_j, -\sin \theta_j); \\ \vec{r}^{(j)} &= (\sin \theta_j \cos \phi_j, \sin \theta_j \sin \phi_j, \cos \theta_j). \end{aligned} \quad (5.5)$$

Once we have parameterized the state space Ω we have to find the invariant integration measure $d\Omega$. We have derived this measure earlier [24] and it reads:

$$d\Omega = \cos^2 \alpha \sin \alpha \sin \theta_1 \sin \theta_2 d\alpha d\psi d\phi_1 d\theta_1 d\phi_2 d\theta_2. \quad (5.6)$$

A set of projectors associated with the observables $\hat{\sigma}^{(1)}$ and $\hat{\sigma}^{(2)}$ [in what follows we use the notation such that the position of the operator to the left (right) of the symbol \otimes is associated with the first (second) spin-1/2]:

$$\hat{P}_{s, \hat{s}_i^{(1)}} = \frac{(\hat{1} + s\hat{\sigma}_i)}{2} \otimes \hat{1}; \quad \hat{P}_{s, \hat{s}_i^{(2)}} = \hat{1} \otimes \frac{(\hat{1} + s\hat{\sigma}_i)}{2}; \quad \hat{P}_{s, \hat{s}_i^{(1)} \hat{s}_j^{(2)}} = \frac{\hat{1} \otimes \hat{1}}{2} + s \frac{\hat{\sigma}_i \otimes \hat{\sigma}_j}{2}. \quad (5.7)$$

The corresponding conditional probabilities do read

$$p(s, \hat{s}_i^{(1)} | \hat{\rho}(\alpha \dots)) = \frac{1}{2} + s \frac{\cos(\alpha)}{2} r_i^{(1)}; \quad p(s, \hat{s}_i^{(2)} | \hat{\rho}(\alpha \dots)) = \frac{1}{2} + s \frac{\cos(\alpha)}{2} r_i^{(2)}; \quad (5.8)$$

$$p(s, \hat{s}_i^{(1)} \hat{s}_j^{(2)} | \hat{\rho}(\alpha \dots)) = \frac{1}{2} + s \frac{r_i^{(1)} r_j^{(2)}}{2} + s \frac{\sin(\alpha)}{2} [(k_i^{(1)} k_j^{(2)} - l_i^{(1)} l_j^{(2)}) \cos \psi - (k_i^{(1)} l_j^{(2)} + l_i^{(1)} k_j^{(2)}) \sin \psi].$$

We remind ourselves that we do consider only measurements performed on the first spin described by the observables $\hat{\sigma}^{(1)}$. After the Bayesian reconstruction of the composed system is performed then the “reservoir” degrees of freedom are traced out. The resulting density operator describes an *a posteriori* estimation of the density operator of a two-level system with no reference on the *a priori* assumption about the purity of the spin.

Instead of analysing estimated density operators after a finite number of measurements performed over the spin-1/2 (i.e. measurements performed over a finite number of elements of the ensemble) we focus our attention on results of the reconstruction in the limit of infinite number of measurements.

A. Observation level $\mathcal{O}_A = \{\hat{s}_z\}$

On the observation level \mathcal{O}_A only the spin component \hat{s}_z is measured. This kind of the measurement can be performed with the help of one Stern-Gerlach apparatus. With the help of the data obtained in a large (infinite) number of measurements we can express the density operator of the spin-1/2 under consideration as (the trace over the reservoir degrees of freedom has already been performed)

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_{-1}^1 y^2 dy \int_0^\pi \sin \theta_1 d\theta_1 \delta(\langle \hat{\sigma}_z^{(1)} \rangle - y \cos \theta_1) (\hat{1} + y \cos \theta_1 \hat{\sigma}_z), \quad (5.9)$$

where the variable α is substituted by $y = \cos \alpha$. When we perform integration over the variable y we obtain the expression

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_{\mathcal{L}'} d\theta_1 \frac{\sin \theta_1}{\cos^2 \theta_1 |\cos \theta_1|} (\hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle \hat{\sigma}_z), \quad (5.10)$$

where the integration is performed over the region

$$\mathcal{L}' := \{0, \pi\} \quad \text{such that} \quad |\cos \theta_1| \geq |\langle \hat{\sigma}_z^{(1)} \rangle|. \quad (5.11)$$

After we perform the integration over θ_1 we obtain for the *a posteriori* estimation of the density operator the expression

$$\hat{\rho} = \frac{1}{2} (\hat{1} + \langle \hat{\sigma}_z \rangle \hat{\sigma}_z). \quad (5.12)$$

which is identical to the one obtained via the Jaynes principle of maximum entropy [28].

B. Observation level $\mathcal{O}_B = \{\hat{s}_z, \hat{s}_x\}$

Let us extend the observation level \mathcal{O}_A and let us assume the measurement of two spin projections \hat{s}_z and \hat{s}_x (i.e. two Stern-Gerlach apparatuses with fixed orientations are employed). In the limit of infinite number of measurements one can express the Bayesian estimation of the density operator of the spin-1/2 on the given observation level as (here the trace over the “reservoir” spin has already been performed):

$$\begin{aligned} \hat{\rho} = \frac{1}{\mathcal{N}} \int_{-1}^1 y^2 dy \int_0^\pi \sin \theta_1 d\theta_1 \int_0^{2\pi} d\phi_1 \delta(\langle \hat{\sigma}_z^{(1)} \rangle - y \cos \theta_1) \delta(\langle \hat{\sigma}_x^{(1)} \rangle - y \sin \theta_1 \cos \phi_1) \\ \times (\hat{1} + y \sin \theta_1 \cos \phi_1 \hat{\sigma}_x + y \sin \theta_1 \sin \phi_1 \hat{\sigma}_y + y \cos \theta_1 \hat{\sigma}_z). \end{aligned} \quad (5.13)$$

When we perform integration over the variable y we find

$$\begin{aligned} \hat{\rho} = \frac{1}{\mathcal{N}} \int_0^{2\pi} d\phi_1 \int_{\mathcal{L}'} d\theta_1 \frac{\sin \theta_1}{\cos^2 \theta_1 |\cos \theta_1|} \delta(\langle \hat{\sigma}_x^{(1)} \rangle - \tan \theta_1 \cos \phi_1 \langle \hat{\sigma}_z^{(1)} \rangle) \\ \times (\hat{1} + \langle \hat{\sigma}_z^{(1)} \rangle \tan \theta_1 \cos \phi_1 \hat{\sigma}_x + \langle \hat{\sigma}_z^{(1)} \rangle \tan \theta_1 \sin \phi_1 \hat{\sigma}_y + \langle \hat{\sigma}_z^{(1)} \rangle \hat{\sigma}_z). \end{aligned} \quad (5.14)$$

The integration over the variable ϕ_1 in the right-hand side of Eq.(5.14) results into the following expression

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_{\mathcal{L}''} d\theta_1 \sum_{j=1}^2 \frac{1}{\cos^2 \theta_1 |\sin \phi_1^{(j)}|} (\hat{1} + \langle \hat{\sigma}_x^{(1)} \rangle \hat{\sigma}_x + \langle \hat{\sigma}_z^{(1)} \rangle \tan \theta_1 \sin \phi_1^{(j)} \hat{\sigma}_y + \langle \hat{\sigma}_z^{(1)} \rangle \hat{\sigma}_z), \quad (5.15)$$

where the integration is performed over the region

$$\mathcal{L}'' := \{0, \pi\} \quad \text{such that} \quad |\cos \theta_1| \geq |\langle \hat{\sigma}_z^{(1)} \rangle|, \quad \text{and} \quad |\tan \theta_1| \geq \left| \frac{\langle \hat{\sigma}_x^{(1)} \rangle}{\langle \hat{\sigma}_z^{(1)} \rangle} \right|. \quad (5.16)$$

The sum in Eq.(5.15) is performed over two values $\phi_1^{(j)}$ of the variable ϕ_1 which are equal to two solutions of the equation

$$\cos \phi_1 = \frac{\langle \hat{\sigma}_x^{(1)} \rangle}{\langle \hat{\sigma}_z^{(1)} \rangle \tan \theta_1}. \quad (5.17)$$

Due to the fact that the term in front of the operator $\hat{\sigma}_y^{(1)}$ is an odd function of $\phi_1^{(j)}$, we can straightforwardly perform integration over θ_1 and we find the expression for the reconstructed density operator

$$\hat{\rho} = \frac{1}{2}(\hat{1} + \langle \hat{\sigma}_z \rangle \hat{\sigma}_z + \langle \hat{\sigma}_x \rangle \hat{\sigma}_x), \quad (5.18)$$

which again is exactly the same as if we perform the reconstruction with the help of the Jaynes principle [28].

C. Complete observation level $\mathcal{O}_C = \{\hat{s}_z^{(1)}, \hat{s}_x^{(1)}, \hat{s}_y^{(1)}\}$

Let us assume now that the measurement is performed with the help of three Stern–Gerlach apparatuses, each of which are measuring three spin component $s_i^{(1)}$ ($i = x, y, z$). On this complete observation level the expression for the Bayesian estimation of the density operator of the spin-1/2 in the limit of infinite number of measurements can be expressed as (here again we have already traced over the “reservoir” degrees of freedoms):

$$\begin{aligned} \hat{\rho} = & \frac{1}{\mathcal{N}} \int_{-1}^1 y^2 dy \int_0^\pi \sin \theta_1 d\theta_1 \int_0^{2\pi} d\phi_1 \delta(\langle \hat{\sigma}_z^{(1)} \rangle - y \cos \theta_1) \delta(\langle \hat{\sigma}_x^{(1)} \rangle - y \sin \theta_1 \cos \phi_1) \\ & \times \delta(\langle \hat{\sigma}_y^{(1)} \rangle - y \sin \theta_1 \sin \phi_1) (\hat{1} + y \sin \theta_1 \cos \phi_1 \hat{\sigma}_x + y \sin \theta_1 \sin \phi_1 \hat{\sigma}_y + y \cos \theta_1 \hat{\sigma}_z). \end{aligned} \quad (5.19)$$

We can rewrite Eq.(5.19) as

$$\begin{aligned} \hat{\rho} = & \frac{1}{\mathcal{N}} \int_{\mathcal{L}''} d\theta_1 \sum_{j=1}^2 \frac{1}{\cos^2 \theta_1 |\sin \phi_1^{(j)}|} \delta \left(\langle \hat{\sigma}_y^{(1)} \rangle - \tan \theta_1 \sin \phi_1^{(j)} \langle \hat{\sigma}_z^{(1)} \rangle \right) \\ & \times \left[\hat{1} + \langle \hat{\sigma}_x^{(1)} \rangle \hat{\sigma}_x + \langle \hat{\sigma}_y^{(1)} \rangle \tan \theta_1 \sin \phi_1^{(j)} \hat{\sigma}_y + \langle \hat{\sigma}_z^{(1)} \rangle \hat{\sigma}_z \right], \end{aligned} \quad (5.20)$$

where \mathcal{L}'' and $\phi_1^{(j)}$ are defined by Eqs.(5.16) and (5.17), respectively. Now the integration over parameter θ_1 can be easily performed and for the density operator of the given spin-1/2 system we find

$$\hat{\rho} = \frac{1}{2}(\hat{1} + \langle \hat{\sigma}_x^{(1)} \rangle \hat{\sigma}_x + \langle \hat{\sigma}_y^{(1)} \rangle \hat{\sigma}_y + \langle \hat{\sigma}_z^{(1)} \rangle \hat{\sigma}_z). \quad (5.21)$$

The density operator (5.21) obtained with the help of Bayesian inference scheme is equal to that one which follows from the Jaynes principle of the maximum entropy. Moreover this same result can be obtained from other reconstruction schemes, such as the discrete quantum tomography [9] (see also [2] and [29]).

The simple example of the spin-1/2 reconstruction reveals deep conceptual relationship between the quantum Bayesian inference and the Jaynes principle of the maximum entropy. To understand this relationship more clearly we turn our attention once again to the *a priori* assumption under which we have performed the reconstruction. We have assumed that the spin-1/2 can be in an impure state (therefore we have applied the “purification” procedure). As a consequence of this assumption the vonNeumann entropy of the reconstructed density operator may be larger than zero, which is equivalent to the fact that the mean values of the observables $\hat{\sigma}_i$ ($i = x, y, z$) do fulfill the condition

$$\langle \hat{\sigma}_x \rangle^2 + \langle \hat{\sigma}_y \rangle^2 + \langle \hat{\sigma}_z \rangle^2 \leq 1, \quad (5.22)$$

i.e. the reconstructed state can be expressed as a point either on or inside the Poincare sphere.

On the contrary, if it is *a priori* assumed that the reconstructed state is a pure one (see, for instance, works by Jones [23]), then the Bayesian reconstruction in the limit of infinite number of measurements on the complete observation level results in the reconstructed density operator which can be expressed as

$$\hat{\rho} = \frac{1}{\mathcal{N}} \int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta \delta(\langle \hat{\sigma}_z \rangle - \cos \theta) \delta(\langle \hat{\sigma}_x \rangle - \sin \theta \cos \phi) \delta(\langle \hat{\sigma}_y \rangle - \sin \theta \sin \phi) \\ \times (\hat{1} + \sin \theta \cos \phi \hat{\sigma}_x + \sin \theta \sin \phi \hat{\sigma}_y + \cos \theta \hat{\sigma}_z). \quad (5.23)$$

The integral in the right-hand side of the equation can only be performed if the mean values of the observables under consideration do fulfill the purity condition

$$\langle \hat{\sigma}_x \rangle^2 + \langle \hat{\sigma}_y \rangle^2 + \langle \hat{\sigma}_z \rangle^2 = 1. \quad (5.24)$$

Providing the purity condition holds then from Eq.(5.23) we obtain for the density operator the expression (5.21), otherwise the reconstruction scheme fails.

The limit formulae for the Bayes inference have an appealing geometrical interpretation. For example the three δ -functions in Eq.(5.23) correspond to three specific orbits on the Poincare sphere, each of which is associated with a set of pure states which posses the measured value of a given observable \hat{s}_i . This corresponds to an averaging over the *generalized microcanonical* ensemble of pure states having the measured mean values of the three observables. The reconstructed density operator then describes a point on the Poincare sphere which coincides with an intersection of these three “orbits”. Consequently, if the three orbits have no intersection, the reconstruction scheme fails, because there does not exist a *pure* state with the given mean values of the measured observables.

On the other hand, the form of the *a posteriori* expression for the reconstructed density operator (5.19) obtained with the help of the Bayesian inference with no *a priori* restriction on the purity of the reconstructed state reveals that the estimated density operator can be obtained as a result of averaging over all points on and inside the Poincare sphere. This averaging over the *generalized canonical ensemble* is equal to the maximization of the vonNeumann entropy as assumed by Jaynes.

VI. CONCLUSIONS

In the paper we have analyzed in detail the logical connection between three different reconstruction schemes: **(1)** If measurements over a finite number of elements of the ensemble are performed then one can obtain the *a posteriori* estimation of the density operator with the help of the Bayesian inference. If nothing is know about the reconstructed state one has to assume a constant prior probability distribution on the parametric state space under the assumption that the system is in a statistical mixture. **(2)** As soon as number of measurements becomes large then the Bayesian inference scheme becomes equal to the reconstruction scheme based on the Jaynes principle of the maximum entropy, i.e., in the limit of infinite number of measurements *a posteriori* estimated density operator fulfills the condition of the maximum entropy. Consequently, it is equal to the generalized canonical density operator. **(3)** If the quorum of observables is measured, then the generalized canonical operator is equal to the “true” density operator of the system itself, i.e. the complete reconstruction via the MaxEnt principle is performed. It is the question of technical convenience which reconstruction scheme on the complete observation level is utilized (for instance, quantum tomography can be used), but the fact is that all of them can be formulated as a maximization of the entropy under given constraints.

Acknowledgements

This work was in part supported by the United Kingdom Engineering and Physical Sciences Research Council, the European Community, and the Grant Agency VEGA of the Slovak Academy of Sciences (grant. n. 2/1154/96). We acknowledge the support by the East-West Program of the Austrian Academy of Sciences under the contract No. 45.367/6-IV/3a/95 of the Österreichisches Bundesministerium für Wissenschaft und Forschung.

-
- [1] L.E. Ballentine: *Quantum Mechanics* (Prentice Hall, Englewood Cliffs, New Jersey, 1990).
 - [2] W. Band and J.L. Park, *Am. J. Phys.* **47**, 188 (1979); *Found. Phys.* **1**, 133 (1970); *Found. Phys.* **1**, 339 (1971); J.L. Park and W. Band, *Found. Phys.* **1**, 211 (1971).
 - [3] J. vonNeumann: *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955) see also J.A. Wheeler and W.H. Zurek: *Quantum Theory and Measurement* (Princeton University Press, Princeton, 1983).

- [4] W. Pauli: *General Principles of Quantum Mechanics* (Springer Verlag, Berlin, 1980); see also B. d’Espagnat: *Conceptual Foundations of Quantum Mechanics*, 2nd ed. (W.A.Benjamin, Reading, 1976).
- [5] W. Gale, E. Guth, and G.T. Trammel, *Phys. Rev.* **165**, 1434 (1968); see also A. Orlowski and H. Paul, *Phys. Rev.* **50**, R921 (1994).
- [6] K.Vogel and H.Risken, *Phys. Rev. A* **40**, 2847 (1989); see also J.Bertrand and P.Bertrand, *Found. Phys.* **17**, 397 (1987); M.Freyberger, K.Vogel, and W.P.Schleich, *Phys. Lett. A* **176**, 41 (1993); H.Kühn, D.-G.Welsch, and W.Vogel, *J. Mod. Opt.* **41**, 1607 (1994); G.S.Agarwal and S.Chaturvedi, *Phys. Rev. A* **49**, R665 (1994); W.Vogel and D.-G.Welsch, *Acta Phys. Slov.* **45**, 313 (1995); U. Leonhardt, M. Munroe, T. Kiss, T. Richter, and M.G. Raymer, *Opt. Commun.* **127**, 144 (1996); U. Leonhardt, H. Paul, and G.M. D’Ariano, *Phys. Rev. A* **52**, 4899 (1995); G.M. D’Ariano, U. Leonhardt, and H. Paul, *Phys. Rev. A* **52**, R1801 (1995); U. Leonhardt and H. Paul, *Phys. Lett. A* **193**, 117 (1994); *J. Mod. Opt.* **41**, 1427 (1994).
- [7] D.T. Smithey, M. Beck, M.G. Raymer and A. Faridani, *Phys. Rev. Lett.* **70**, 1244 (1993); D.T. Smithey, M. Beck, J. Cooper and M.G. Raymer, *Phys. Scr. T* **48**, 35 (1993); M. Beck, M.G. Raymer, I.A. Wamsley, and V. Wong, *Opt. Lett.* **18**, 2041 (1993). M. Beck, D.T. Smithey, and M.G. Raymer, *Phys. Rev. A* **48**, R890 (1993); see also M.G. Raymer, M. Beck, and D.F. McAlister, *Phys. Rev. Lett.* **72**, 1137 (1994); M.G. Raymer, D.T. Smithey, M. Beck, and J. Cooper, *Acta Phys. Pol.* **86**, 71 (1994).
- [8] U. Janicke and M. Wilkens, *J. Mod. Opt.* **42**, 2183 (1995).
- [9] U.Leonhardt, *Phys.Rev.Lett.* **74**, 4101 (1995); *Phys. Rev. A* **53**, 2998 (1996); see also W.K.Wootters, *Found. Phys.* **16**, 391 (1986). W.K.Wootters, *Ann. Phys. (N.Y.)* **175**, 1 (1987); D.Galetti and A.F.R. De Toledo Piza, *Physica A* **149**, 267 (1988).
- [10] K. Husimi, *Proc. Phys. Math. Soc. Jpn.* **22**, 264 (1940); Y. Kano, *J. Math. Phys.* **6**, 1913 (1965).
- [11] E. Arthurs and J.L. Kelly, Jr., *Bell. Syst. Tech. J.* **44**, 725 (1965); W.K. Wootters and W.H. Zurek, *Phys. Rev. D* **19**, 473 (1979); Y.Lai and H.A.Haus, *Quantum Opt.* **1**, 99 (1989); D.Lalović, D.M.Davidović, and N.Bijedić, *Phys. Rev. A* **46**, 1206 (1992); D.M.Davidović and D.Lalović, *J. Phys. A* **26**, 5099 (1993); S.Chaturvedi, G.S.Agarwal, and V.Srinivasan, *J. Phys.A* **27**, L39 (1994); M.G.Raymer, *Am. J. Phys.* **62**, 986 (1994); U. Leonhardt and H. Paul, *Phys. Rev. A* **48**, 4598 (1993); V. Bužek, C.H. Keitel, and P.L. Knight, *Phys. Rev. A* **51**, 2575 (1995); see also the review articles S.Stenholm, *Ann. Phys. (N.Y.)* **218**, 233 (1992), and U. Leonhardt and H. Paul, *Quant. Electron.* **19**, 89 (1995).
- [12] K. Wódkiewicz *Phys. Rev. Lett.* **52**, 1064 (1984); *Phys. Lett. A* **115**, 304 (1986); *Phys. Lett. A* **129**, 1 (1988).
- [13] N.G. Walker and J.E. Carroll, *Opt. Quant. Electron.* **18**, 335 (1986); N.G. Walker, *J. Mod. Opt.* **34**, 15 (1987).
- [14] J.W. Noh, A. Fougères, and L. Mandel, *Phys. Rev. Lett.* **67**, 1426 (1991); *Phys. Rev. A* **45**, 424 (1992).
- [15] J. Peřina, Z. Hradil, and B. Jurčo: *Quantum Optics and Fundamentals of Physics* (Kluwer Academic Publishers, Dordrecht, 1994).
- [16] T. Opatrný, V. Bužek, J. Bajer, and G. Drobný, *Phys. Rev. A* **52**, 2419 (1995); T. Opatrný, D.-G. Welsch, and V. Bužek, *Phys. Rev. A* **53**, 3822 (1996).
- [17] E.T. Jaynes, *Phys. Rev.* **108**, 171 (1957); *ibid.* **108**, 620 (1957) 620; *Am. J. Phys.* **31**, 66 (1963). E.T. Jaynes, “Information theory and statistical mechanics”, in *1962 Brandeis Lectures, Vol 3*, ed. K.W. Ford (Benjamin, Inc. New York, 1963), p.181. See also E. Fick and G. Sauer mann: *The Quantum Statistics of Dynamic Processes* (Springer Verlag, Berlin, 1990); J.N.Kapur and H.K.Kesavan: *Entropic Optimization Principles with Applications* (Academic Press, New York, 1992).
- [18] The *MaxEnt* principle has a very close relation to the Laplace principle of indifference [see, for instance, H. Jeffreys: *Theory of Probability* (Oxford Univ. Press., Oxford, 1960)] which states that where nothing is known one should choose a constant-valued function to reflect this ignorance. This obviously maximizes any uncertainty measure.
- [19] V. Bužek, G. Adam, and G. Drobný, *Ann. Phys. (N.Y.)* **245**, 37 (1996).
- [20] H. Wiedemann: *Quantum tomography with the maximum entropy principle* (unpublished).
- [21] C.W. Helstrom: *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
- [22] A.S. Holevo: *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
- [23] K.R.W. Jones, *Ann. Phys. (N.Y.)* **207**, 140 (1991); *Phys. Rev. A* **50**, 3682 (1994). For more discussion see Z. Hradil: *Quantum state estimation*, Los Alamos e-print archive, quant-ph/9609012.
- [24] R. Derka, V. Bužek, and G. Adam, *Acta Phys. Slov.* **46**, 355 (1996); see also R. Derka, V. Bužek, G. Adam, and P.L. Knight (unpublished).
- [25] I.S. Gradshteyn and I.M. Ryzhik: *Table of Integrals, Series, and Products* (Academic Press, New York, 1980).
- [26] A. Uhlmann, *Rep. Math. Phys.* **9**, 273 (1976); *ibid* **24**, 229 (1986).
- [27] A.K. Ekert and P.L. Knight, *Am. J. Phys.* **63**, 415 (1995).
- [28] G. Drobný, R. Derka, G. Adam, and V. Bužek: “Reconstruction of quantum states of spin systems via the Jaynes principle of maximum entropy” to appear in the special issue of *J. Mod. Opt.* (1997).
- [29] R.G.Newton and Bing-Lin Young, *Ann. Phys. (N.Y.)* **49**, 393 (1968).